DOCUMENT RESUME

ED 465 344 HE 034 994

AUTHOR Schaeffer, Evonne L.; McGrady, Jennifer A.; Bhargava, Tina; Engel, Claudia

TITLE Online Debate To Encourage Peer Interactions in the Large Lecture Setting: Coding and Analysis of Forum Activity.

PUB DATE 2002-04-00

NOTE 24p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 1-5, 2002). Small type in some figures and colors in others may not reproduce well.

PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS *Biology; *Computer Mediated Communication; Group Discussion; Higher Education; Online Systems; *Policy Formation; *Public Policy; *Undergraduate Students; *World Wide Web

IDENTIFIERS Forums; Stanford University CA

ABSTRACT

The culminating course in the Human Biology core sequence at Stanford University, California, is a course that considers the relationship between the social and natural sciences and the development of wise public policy. A Web-based forum was developed to facilitate weekly debate on policy challenges presented in this course, and this forum was evaluated by coding and analyzing student postings from each of 2 sections of 21 students. Over the quarter, students participated in seven policy challenge assignments. Although using the forum was not an assignment with which students were comfortable immediately, they learned to appreciate the opportunity to challenge each other in an online environment, and they learned to build on each others posts and to participate at more than the required amount. Over time, many discussions became more sustained, more focused, and better elaborated. Some students postings did not improve over time, and the increased average quality of postings was only significant for one section. The implications of these findings and the need for additional studies are discussed. (Contains 5 figures and 21 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 465 344

Online Debate to Encourage Peer Interactions in the Large Lecture Setting: Coding and Analysis of Forum Activity

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

E Schaeffer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Evonne L. Schaeffer[1]
Jennifer A. McGrady
Tina Bhargava
Claudia Engel

Stanford University

Paper presented at the annual meeting of the American Educational Research Association. April 1 – 5, New Orleans, Louisiana.

[1] Email correspondence regarding this paper should be addressed to Evonne Schaeffer, evonne@stanford.edu.

HE034994

2

BEST COPY AVAILABLE

Background

Large lecture courses, which dominate many students' undergraduate experiences, typically follow the transmission model of teaching and learning with the students sitting as passive recipients of the instructor's knowledge (Boyer Commission, 1998; Friedlander & Kerns, 1998; Twigg, 1999). Student interactions are often not encouraged, despite ample research demonstrating the cognitive and motivational benefits of students interacting with one another (e.g., Bruffee, 1999; Roschelle, 1992). Computer-mediated communications provide rich possibilities in such situations (Hara, Bonk, & Angeli, 2000; Jermann & Dillenbourg,1999; Pena-Shaff, Martin, & Gay, 2001; Veerman, Andriessen, & Kanselaar,1999).

Members of the Stanford Learning Lab worked closely with Professor Donald Kennedy to develop, implement, and evaluate an online debate assignment that allows groups of students the opportunity to argue a policy position with their peers. We implemented this assignment for two years in a large lecture course on health and environmental policy-related issues.

Human Biology is one of the largest undergraduate majors at Stanford University. As the culminating course in the required Human Biology sophomore core sequence, Human Biology 4B is defined by the department as a course that encourages students to consider "the relationship between the social and natural sciences—the disciplines of Human Biology—and the development of wise public policies." Course assignments are intended to encourage students to apply material from previous courses to construct solutions to authentic problems. This course typically has 250 students; it is taught with three to four lectures plus one small (n=20 students) group discussion section each week.

The Intervention: Online Policy Challenges

A web-based forum was developed to facilitate weekly debate on the policy challenges. The online debate forum was designed to enhance peer-to-peer learning, encourage thought on the policy issues themselves, and to improve students' ability to make persuasive arguments. Not simply a vehicle to discuss the course readings, the online debate forum provided students with opportunities to actively engage in the content by arguing for a certain policy position and attempting to sway their peers to their viewpoint (Engel & Schaeffer, 2001). In this course, persuasive argumentation became both the means and the ends to learning (Lunsford & Ruszkiewicz, 1999; Voss & Means, 1991). Student participation in the online forum discussions was required and graded.

The Weekly Process

This assignment was first implemented and evaluated in the Spring 1999. Based on the evaluation results (Schaeffer et al., 1999), modifications were made for the Spring 2000 implementation. Each section of approximately 20 students had a separate online debate forum. Each week, students were presented with a topic and two courses of action from which to choose (e.g., competing policy options for reducing the population growth rate in a developing country). Each student took an initial stand on the issue and provided a justification for their position. Pairs of students were assigned as "lead analysts" each week to synthesize the array of opening rationales. These syntheses served as the beginning of the online forum debate among members of each section.

Students debated the issues with their peers for 2 1/2 days. Figure 1 shows an excerpt of such a debate. After the debating period was over, students would again vote for a policy option, perhaps shifting their opinion.

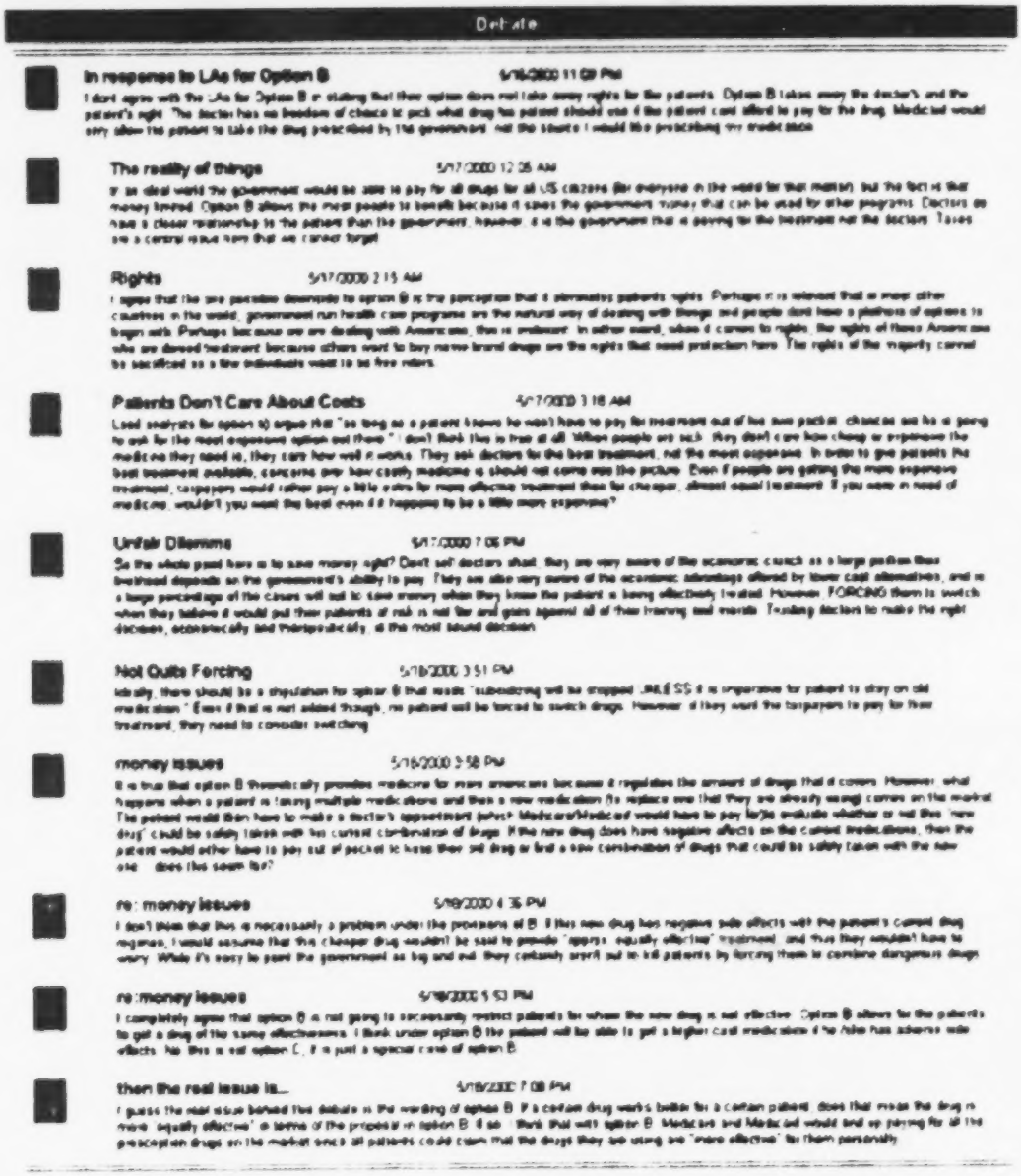
Debate
In response to LAs for Option B
I dont agree with the LAs for Option B in stating that their option does not take away rights for the patients. Option B takes away the doctor's and the patient's right. The doctor has no freedom of choice to pick what drug his patient should use if the patient cant afford to pay for the drug. Medicaid would only allow the patient to take the drug prescribed by the government, not the doctor. I would like prescribing my medication
The reality of things
If an ideal world the government would be able to pay for all drugs for all US citizens (for everyone in the world for that matter), but the fact is that money is limited. Option B allows the most people to benefit because it saves the government money that can be used by other programs. Doctors do have a closer relationship to the patient than the government, however, it is the government that is paying for the treatment not the doctors. Taxes are a central issue here that we cannot forget
Rights
I agree that the one possible downside to option B is the perception that it eliminates patients rights. Perhaps it is relevant that in many other countries in the world, government run health care programs are the natural way of dealing with things and people dont have a plethora of options to begin with. Perhaps because we are dealing with Americans, this is irrelevant. In other words, when it comes to rights, the rights of those Americans who are denied treatment because others want to buy name brand drugs are the rights that need protection here. The rights of the majority cannot be sacrificed so a few individuals want to be free riders.
Patients Don't Care About Costs
Lead analysts for option a) argue that "as long as a patient knows he won't have to pay for treatment out of his own pocket, chances are he is going to ask for the most expensive option out there." I don't think this is true at all. When people are sick they don't care how cheap or expensive the medicine they need is, they care how well it works. They ask doctors for the best treatment, not the most expensive. In order to give patients the best treatment available, concerns over how costly medicine is should not come into the picture. Even if people are getting the more expensive treatment, taxpayers would rather pay a little extra for more effective treatment than for cheaper, almost equal treatment. If you were in need of medicine, wouldn't you want the best even if it happens to be a little more expensive?
Unfair Dilemma
So the whole point here is to save money right? Dont tell doctors what they are very aware of the economic crunch as a large portion their livelihood depends on the government's ability to pay. They are also very aware of the economic advantages offered by lower cost alternatives, and in a large percentage of the cases will act to save money when they know the patient is being effectively treated. However, FORCING them to switch when they believe it would put their patients at risk is not fair and goes against all of their training and morals. Trusting doctors to make the right decision, economically and therapeutically, is the most sound decision.
Not Quite Forcing
Ideally, there should be a stipulation for option B that reads "subsidizing will be stopped UNLESS it is imperative for patient to stay on old medication." Even if that is not added though, no patient will be forced to switch drugs. However, if they want the taxpayers to pay for their treatment, they need to consider switching.
money issues
It is true that option B theoretically provides medicine for more americans because it regulates the amount of drugs that it covers. However, what happens when a patient is taking multiple medications and then a new medication is replace one that they are already using comes on the market. The patient would then have to make a doctor's appointment (since Medicare/Medicaid would have to pay for)to evaluate whether or not this "new drug" could be safely taken with his current combination of drugs. If the new drug does have negative affects on the current medications, then the patient would either have to pay out of pocket to keep their old drug or find a new combination of drugs that could be safely taken with the new one. does this seem fair?
re: money issues
I don't think that this is necessarily a problem under the provisions of B. If this new drug has negative side effects with the patient's current drug regimen, I would assume that this cheaper drug wouldn't be said to provide "approx. equally effective" treatment, and thus they wouldn't have to worry. While it's easy to paint the government as big and evil, they certainly aren't out to kill patients by forcing them to combine dangerous drugs.
re:money issues
I completely agree that option B is not going to unreasonably restrict patients for whom the new drug is not effective. Option B allows for the patients to get a drug of the same effectiveness. I think under option B the patient will be able to get a higher cost medication if he/she has adverse side effects. So this is not option C, it is just a special case of option B.
then the real issue is...
I guess the real issue behind the debate in the wording of option B. If a certain drug works better for a certain patient, does that mean the drug is more "equally effective" in terms of the proposal in option B. If so, I think that with option B, Medicare and Medicaid would end up paying for all the prescription drugs on the market since all patients could claim that the drugs they are using are "more effective" for them personally.

Figure 1. Excerpt from forum from Week 8: MedicAid reimbursement. Section 2; Thread 3.

ENTIRE DOCUMENT:
POOR PRINT QUALITY

BEST COPY AVAILABLE

Goals for the Activity

The course materials (handout entitled "Policy Challenges Overview") stated, and the teaching staff emphasized, the following purposes of the policy challenge assignment:

- "hone (your) debating skills and focus upon composing persuasive arguments and synthetic statements"
- "synthesize information, reason from known facts, analyze data, and develop thoughtful conclusions"
- "make intelligent connections with material from lecture and readings"
- "take into account your classmates' postings and use them to help further debate"
- "try wherever possible to incorporate new or creative perspectives in your comments"

It should be noted that although practicing and improving argumentation were part of the goals of this activity, students were not provided any explicit formal instruction in debating or in argumentation, which was consistent with the "mood" of the course as a whole in emphasizing peer-to-peer learning.

Evaluating the Online Debate Activity

We evaluated the online debate activity from various perspectives, including observing students interacting with it and surveying students' opinions. In addition, we developed codes to systematically categorize the activity on the debate forum. The purpose of these evaluation efforts was not so much to assess individual students; rather to understand how students experience the assignment and whether the design of the assignment encouraged practice of, and therefore improvement in, argumentation skills.

Background to Developing Codes

In aligning with the instructor's emphasis of the importance of argumentation, we explored models of argumentation and persuasion as to how they could inform our coding process. A pilot test of a coding scheme based on Toulmin's model (evaluating elements of arguments: claim, warrant and data; backing, rebuttal and qualifier) was unsuccessful because it was impracticable to distinguish the key elements within one posting (Van Eemerean et al., 1987). This approach does not address the asynchronicity and interactivity that define the online debate forum. As Henri (1992) states, "We cannot analyze a CMC text as we might a 'constructed' text ... a participant's contributions must be considered both singly and in relation to those of the others" (p. 119).

A number of studies have proposed methodologies for the analysis of online discourse. Henri (1992) proposed a model to "highlight five dimensions of the learning process exteriorized in messages: participation, interaction, social, cognitive, and metacognitive dimensions" (p. 117). While Henri's model is widely recognized as providing an initial framework, it has been criticized for the lack of clarity of its proposed coding categories. According to Hara et al. (2000), Henri's model "lacks detailed criteria for systematic and robust classification of electronic discourse" (p. 118). Additionally, Howell-Richardson and Mellar (1996) criticized Henri's "unit of meaning" for being "undefined" (p. 51). Furthermore, despite her call for a system that enables educators to "interpret messages rapidly," Henri's model is detailed, complex, and time-consuming to implement (p. 121).

Building on Henri's model, Howell-Richardson and Mellar (1996) proposed the creation of message maps to "establish the extent to which...activity is distributed throughout the group" (p. 53). More recent work by Hara et al. (2000) and Pena-Shaff et al. (2001) developed message maps to represent the interactions of postings in graphic form. These studies, though more aligned with analyzing tools for general discussion as compared to argumentative discourse, provide work complementary to ours.

Most aligned with our work are studies that deal directly with analyzing online argumentation. Veerman, Andriessen, and Kanselaar (1999) characterized students' dialogues "in terms of their constructive and argumentative contributions," (p. 1) and captured this aspect of students' postings by recording the categories of "argumentative information exchanges" (checks, challenges, and counters) that were evidenced in students' postings. Martunnen (1998) evaluated the quality of students' emails in a study where students "practiced argumentation by electronic mail" (p. 387). Several variables were coded, including interactivity, the references to other students' arguments within each student's emails, and the direction (agree/disagree) and quality (grounded/non-grounded) of each reference (pp. 392-393).

Using a different approach, Jermann and Dillenbourg (1999) analyzed the arguments of students in the "Argue Graph" learning activity, focusing on "the effects of discussion and opinion conflict on the elaboration of argument" (p. 1). Their Argue Graph activity resembled the policy challenge assignment:

students first had to commit to one option in a theoretical question and give a rationale for their choice, then had to debate with a student who had committed to another option and reach agreement on which option to select, giving a shared rationale for this choice. Jermann and Dillenbourg coded students' arguments into one of the following: justification (the argument contained "at least one idea not present in the phrasing of the selected option") or reformulation (the argument did not "contain an idea not present in the selected option") (p. 5). Their analysis revealed that students' arguments tended to include more "justification" than "reformulation" in the shared than in the individual rationales, perhaps because as the authors note, the design of the activity led students to "make (their argument) explicit and to elaborate it" (p. 8).

Our coding scheme draws on those used by Veerman et al., Jermann and Dillenbourg, and Martunnen. Various elements of their research seemed appropriate because: 1) these studies focused on evaluation of students' argumentation, and not just general discussion; 2) the design and purpose of the activities these studies evaluated were similar to goals of the policy challenge assignment; and 3) the criteria used to assess students' arguments were consistent with the policy challenge instructions given to students to guide their online debate postings. Most importantly, these studies emphasized peer-to-peer interactions.

Method

To address the increase in Type I errors that can result from developing and analyzing codes on the same sample, we developed codes on one section and conducted the analyses on two other sections from the Spring 2000 implementation. Analyses are conducted within each section to see whether trends observed in one section replicate in the other.

The sections were chosen to minimize potential confounds: they were led by different course assistants, each had 21 students, and met at the same time each week. Gender was similarly represented in the two sections; Section 1 had 14 females and 7 males; Section 2 had 13 females and 8 males. These two sections do not differ from the class as a whole in terms of instructor-assigned grades (e.g., policy challenges, midterm, final, participation, and course grade) or the two learning style variables, Comfort with Ambiguity and Comfort in Learning from Peers.

Over the course of the quarter, students participated in seven policy challenge assignments. We sampled the second and sixth ones to study, to enable a contrast between early and late in the quarter and to avoid the atypical first and last weeks. The first topic was concerned with population control policy in a developing country; the second topic dealt with criteria for reimbursement of Medicaid funds for new medications and procedures. Both topics had the same level of controversy, with students in each section equally split between the two policy options provided. Table 1 shows the text of the topics and choices for the two weeks we analyzed.

BEST COPY AVAILABLE

Table 1.
Policy Challenge Text

Policy Challenge #2—Week 2

You are the Minister of Family Planning for Ruritania, a developing country with an annual population growth rate of 2% per year. 45% of the population is age 14 and under. Your Ministry has been asked by the President to develop a population control policy; he reminds you that he is capable of imposing fairly stringent regulations on Ruritania's citizens, and he wants to see dramatic reductions in growth rate.

Option A. Launch a massive nation-wide publicity campaign advertising the value of the "one-child family," supported by extensive improvements in the family planning services provided by the Ministry of Health.

Option B. Make investments to improve Ruritania's educational system, and eliminate dependent tax benefits for women who have children before age 25.

Policy Challenge #6—Week 8

You serve as chair of the State Reimbursement Commission, charged with evaluating whether the state should provide out of MedicAid funds payment for new medications and procedures. The first task is to develop criteria for reimbursement; a subcommittee has presented two proposals, and cannot decide which to send forward to the full commission. which do you choose to support?

Option A. Published data on clinical trials (if a drug) or outcome evaluations (if a procedure, plus approval of the relevant federal regulatory body and the appropriate professional academy (e.g., American College of Obstetrics, etc.)

Option B. Published data on clinical trials (if a drug) or outcome evaluations (if a procedure), plus a provision that if a lower-cost alternative of approximately equal effectiveness is available the reimbursement request should be denied.

Unit of Analysis

Each student made several postings on each of the policy challenges, and many postings contain more than one discrete point. This inquiry draws on five units of analysis: section, policy challenge, student, posting, and point.

Developing Codes for Types of Exchanges

Consistent with the goals of the assignment we focused the coding of student postings on peer-to-peer interaction. We created a coding variable, "Type of Exchange" to capture the nature of student interactions in the online forum. It was adapted from the three coding schemes previously described that were largely consistent with the goals of the assignment and the instructions that the students received: Veerman, Andriessen, and Kanselaar's (1999) "Categories of Information Exchange," Jermann & Dillenbourg's (1999) "Degree of Elaboration, " and Martunnen's (1998) interaction analysis and assessment of "grounding." The interrelatedness of the postings were coded at the point level within each forum within each section.

"Type of Exchange" indicates whether a student's point was related to a previous post, and, if so, whether it was agreeing or opposing that post. It also indicates whether the point introduced a new element or simply revisited old ideas, relative to the points that preceded it in the forum. Coding that captures freshness of information in a posting, and how it relates to postings of other students, is consistent with instructions given to students on how to approach this assignment and on how it would be graded.

Table 2 shows the definitions of the five-point *Type of Exchange* variable, with examples. Note that the coding of Type of Exchange depends on the postings earlier in the same debate forum (not shown here) to determine whether a *new* element was introduced.

Table 2.
Definitions and Examples of Type of Exchange Categories

-2 (Counter)	-1 (Challenge)	0	+1 (Acceptance)	+2(Enhancement)
DISAGREE		Unrelated	AGREE	

Counter (coded -2). Applies to those points that, implicitly or explicitly, oppose an earlier posting and introduce a new element of information, which either refers to new facts (e.g., synthesizing course material), or to a justification or elaboration of the points made, or to a new take on facts that have already been introduced to the debate (e.g., providing an emotional appeal based on a factual point somebody already made).

"How is an advertising campaign guaranteed to produce 'dramatic and immediate' changes? Ads are not enforceable policies; therefore, it is too idealistic to expect that ads championing a one child family will be effective."

"Even if contraceptives and family planning services are widely available, I think that without the basis of the educational system in Ruritania strengthened, the children of these families will never be able to break out of the cycle of poverty."

Challenge (coded –1). Applies to those points that, implicitly or explicitly, oppose an earlier posting, but without introducing any new element.

"The elimination of tax benefits for women under 25 presents a tangible impetus for women to avoid having children earlier. I don't agree that this will just result in having a cohort of women who can't support their children. The elimination of tax benefits is far sufficient enough to prevent this from happening in the first place."

"Option B is not 'wasting time' at all, in fact, it is providing a logical plan for the future by addressing the population of the present."

Unrelated (coded 0) applies to those points that make no clear reference, explicitly or implicitly, to the posting of another student.

*"Although both options are *technically* feasible, the more compelling option, in my view, is B."*

"In reviewing all the debate thus far, I was and remain a strong supporter of B."

"Changing culturally determined fertility aspiration levels can be initiated in a non-threatening, respectful way by integrating the value of small families into the popular media. Studies now show that people are generally aware of the costs of children; it is time to turn from educating them about costs they are already aware of to influencing the more deeply ingrained cultural context in which individuals make reproductive decisions."

Acceptance (coded +1) applies to those points that, implicitly or explicitly, support an earlier posting, but without introducing any new element.

"David brings up a good point. Access does not necessarily mean acceptance."

"I totally agree with Lisa. Education reform, while a wonderful and equitable idea, is an extremely long process. Increased family planning and sexual health education as well as increased availability of contraceptives will help keep the population under control until more long term development can be undergone."

Enhancement (coded +2) applies to those points that, implicitly or explicitly, support an earlier posting and introduce a new element.

"I agree with Steve. By implementing wide distribution of free contraception, we are merely ensuring that an unwanted pregnancy would not result from "actions" that kids are just going to do anyway. Even well-educated teenagers in developed countries are sexually active."

"I think Caroline brings up a good point with her Bongaarts quote. Gender equity is extremely important. As we learned in the fall, fertility is lower if women have similar educational opportunities to men. If we assume that the policy in Option B will improve the educational system for both men and women (especially women), then this plan can help in four ways: 1. increased knowledge spread via written word 2. increased knowledge of health issues 3. increased opportunities outside the home, and 4. increased status (especially for women)."

Variables Coded

Each post had several identifying and coded variables: thread number; post order within thread; date/time of posting; length of post measured by number of characters; position of policy challenge supported (Option A, Option B, neutral), and number of points. Postings contain more than one point when they have either more than one content area or more than one distinct idea on the same content. Some postings contained three or four points, and were considered less focused than postings that were limited to

BEST COPY AVAILABLE

one or two points. Each point was coded on Type of Exchange in the context of the section in which it occurred, and topic area.

Aggregates of these variables were then formed at the student level within each section and within each policy challenge, including: number of posts; number of exchanges; number of agreement and disagreement exchanges; average absolute value of exchanges (collapsing 2 and –2; and 1 and –1; etc.) within students across postings; number of times each student's posts were referred to by others, etc. Student-level data were also available for course-grades (e.g., midterm exam, final exam, policy challenge assignment, and participation grades) and survey results (including two short learning styles scales).

Interrater Agreement

Two independent coders worked to refine the coding categories for each policy challenge. A third section of 20 students was used in this preliminary phase. Once agreement was established that exceeded 80% interrater agreement, the two coders worked independently on the two forums within the two different sections. They performed spot checks of each other's sections, and maintained at least 85% interrater agreement throughout the coding, on both the topic and Type of Exchange categories. Instances of disagreement were resolved by discussing and then coming to consensus on the codes.

Results

Overall Summaries

Table 3 shows the overall summaries for the four forums we coded. For both sections, the number of postings to the online forum decreased from early in the quarter to later, from 60 to 50 in Section 1 and from 87 to 70 in Section 2. Although the average number of postings per student decreased over the quarter, from 2.9 to 2.4 in Section 1 and from 4.2 to 3.3 in Section 2, every student with only one exception continued to post at least the required amount (2) for each week. The number of points showed a similar decrease from early to late in the quarter, as did the average number of points per posting, perhaps suggesting that students' individual postings became more focused over time.

Most dramatically, the number of threads (defined as a new posting, followed by a series of replies) declined from the early to later in the quarter, from 26 to 11 in Section 1 and 22 to 8 for Section 2. This decline in the number of threads was accompanied by an increase in the average number of postings per thread, from 2.3 to 4.5 for Section 1 and 4.0 to 8.7 in Section 2. This increase in thread lengths suggests that more sustained discussions occurred later in the quarter compared to earlier.

Although overall there were fewer posts later in the quarter, the evidence supports that those posts were generally better formulated. The percent of unrelated points, not connecting in any obvious way to previous points, went from 27% down to 13% in Section 1 and from 35% down to 26% in Section 2. Additionally, although the percent of points that elaborated on previous points stayed about the same in Section 1 (51% to 52%) it increased in Section 2 (from 35% to 49%). This suggests that overall, students showed improvements in elaborating their ideas and building on the ideas of their peers.

Table 3.
Description of Online Forum Posting Activity Codes

	Section 1		Section 2	
	Week 2	Week 8	Week 2	Week 8
Topic	Population	Medicaid	Population	Medicaid
Number of postings	60	50	87	70
Number of points	115	80	133	92
Number of threads	26	11	22	8
Average postings/student	2.9	2.4	4.2	3.3
Range postings/student	2 – 4	1 – 4	2 – 12	2 - 7
Average points/posting	1.9	1.6	1.5	1.3
Average words/posting	140	116	118	100
Percent posts referencing previous posts	71%	88%	58%	76%
Average postings/thread	2.3	4.5	4.0	8.7

	Section 1		Section 2	
	Week 2	Week 8	Week 2	Week 8
Type of Exchanges[a]				
Counter (-2)	38%	31%	27%	38%
Challenge (-1)	5%	13%	17%	11%
Unrelated (0)	27%	13%	35%	26%
Acceptance (+1)	17%	20%	13%	14%
Enhancement (+2)	12%	21%	7%	11%
Combined Exchanges[b]				
Counter and Enhancement (± 2)	50%	52%	34%	49%
Challenge and Acceptance (± 1)	22%	33%	30%	25%
Unrelated (0)	27%	13%	35%	26%

[a] p = .02 and .25, for Chisquare test for changes from Week 2 to 8 for Sections 1 and 2, respectively.
[b] p = .03 and .11, for Chisquare test for changes from Week 2 to 8 for Sections 1 and 2, respectively.

BEST COPY AVAILABLE

Student-Level Analyses

To capture the extent that each student elaborated on previous peer postings, we calculated the average absolute value for each student's set of Type of Exchange scores. Higher levels of |*Exchange*| indicate better quality posts, that is, ones that elaborate on previous posts and add a new element. Table 4 shows the correlations between the quality of the postings (as measured by the average |*Exchange*|) with student background measures and other measures of course success.

First, we note that quality of posting is not correlated with sex of poster, verbal ability of students, or length of posting (i.e., it is *not* the case that enhancement postings are longer than acceptance postings). Similarly, quality of postings are not correlated with learning styles scales which measure how comfortable students are in learning material without a single correct answer or with learning from peers, except in one case in Week 2 where higher quality posts were generated by students who were less comfortable learning from their peers. One interpretation of these results is that the design of the activity supported students equally across gender, verbal ability, and learning styles. Perhaps it enabled even those students who said they were uncomfortable in ambiguous learning environments to participate successfully. On the other hand, quality of postings were correlated with policy challenge scores assigned by the instructors in both sections, and with other measures of course success for Section 2, presenting a source of evidence for reliability and convergent validity for the Type of Exchange codes.

Table 4.
Correlations of Quality of Postings (|Exchange|) with Background Variables and Course Grades

	Section 1		Section 2	
	Week 2	Week 8	Week 2	Week 8
Number of students	21	21	21	21
Sex	.05	-.05	.16	.39
Length of post	-.00	-.13	.18	.35
Verbal ability self-ranking	-.04	-.26	-.08	.30
Learning styles				
Comfort with ambiguity	.42	.36	-.24	.23
Comfort w/ peer learning	.00	-.23	-.62*	-.08
Course grades				
Policy challenge	.54*	.37	.45*	.42
Midterm	.05	-.14	.52*	.32
Final	.09	-.06	.50*	.65*
Curve (overall score)	.29	-.04	.56*	.62*

* $p < .05$ two-tailed significance.

To address whether students improved in their argumentation (as measured by the Type of Exchange code) we conducted within-person comparisons, using paired t-tests. Whereas the average |Exchange| increased in both sections (from 1.2 to 1.4 in Section 1 and from 1.0 to 1.3 in Section 2), only that change in Section 2 was statistically significant ($p = .20$ and .03, respectively). In each section, 8 students' average |Exchange| decreased over time and twelve students improved, with one student in each section showing no change.

Displaying Section-Level Activity

Building on the graphic representations as used by Hara et al. (2000) and Pena-Shaff et al. (2001), among others, we constructed activity graphs to represent the debate forum activity. The activity graphs display the flow of the forum as a function of time. Much information is displayed in the activity graphs by employing a variety of shapes, colors, and line thicknesses. Of particular interest here is the nature of the interactivity on the forum: which posts refer back to previous posts and how.

Figures 2 to 5 present the activity graphs for Sections 1 and 2 on Weeks 2 and 8. The vertical axis represents the timeline of the assignment; the threads unfold on the horizontal axis. The shape of the points indicate which option the point is supporting: squares represent support for Option A; circles represent support for Option B; and triangles represent posts that are neutral. Posts with multiple points are displayed as multiple adjacent shapes (i.e., square, circle, or triangle). The types of lines represent how postings refer back to previous postings: dashed lines indicate agreement or a challenge and solid lines indicate an enhancement or a counter. The red lines indicate disagreement; the green lines indicate agreement.

Each activity graph displays the overall flow of the forum, with its sequential starts of new threads, times that the densities of the postings are lower (e.g., the first day of the assignment for Section 1, policy challenge 2) and higher (e.g., near the due time in Section 1, policy challenge 6), and types and frequencies of peer-to-peer interactions, both within and across threads. Unrelated posts do not have lines connecting them to previous posts and clearly appear isolated. Some threads show much activity (e.g., Section 2, policy challenge 2, thread 15) while others show sparse activity (e.g., Section 2, policy challenge 2, thread 20). Some postings spark much comment, as represented by many lines connecting back to them (e.g., Section 2, policy challenge 6, thread 5, 4 PM).

However, one prominent visual feature—the line lengths—may be highlighting somewhat irrelevant information. Within a thread, the line length indicates the amount of time elapsed between a posting and its referent; however, across threads (horizontally), the length of the line is meaningful only to the extent that the posting refers back to a posting in a different thread. More often than not these cross-thread referrals occur as the first post of a new thread, which raises the question as to how the students understood the meaning of replying to an existing thread versus starting a new one.

The activity graphs particularly highlight changes in the nature of the debate when comparing the forum activity early in the quarter with that later in the quarter. Over time, as the graphs for both sections show, the discussion became more focused (fewer threads and fewer points per post), more sustained (most of the threads are active for at least 24 hours), better elaborated (more posts per thread), and more interconnected (fewer unrelated points). *That* these changes occurred over time is evident; *why* they occurred is less clear. Perhaps students were simply conserving effort as the end of the quarter drew near; alternatively, they may be demonstrating higher quality posting behavior. Apart from the feedback that they routinely received from their teaching assistants on their posting behavior, the students did not experience any "intervention." These activity graphs were constructed after the quarter was over and did not influence the students' posting behavior.

BEST COPY AVAILABLE

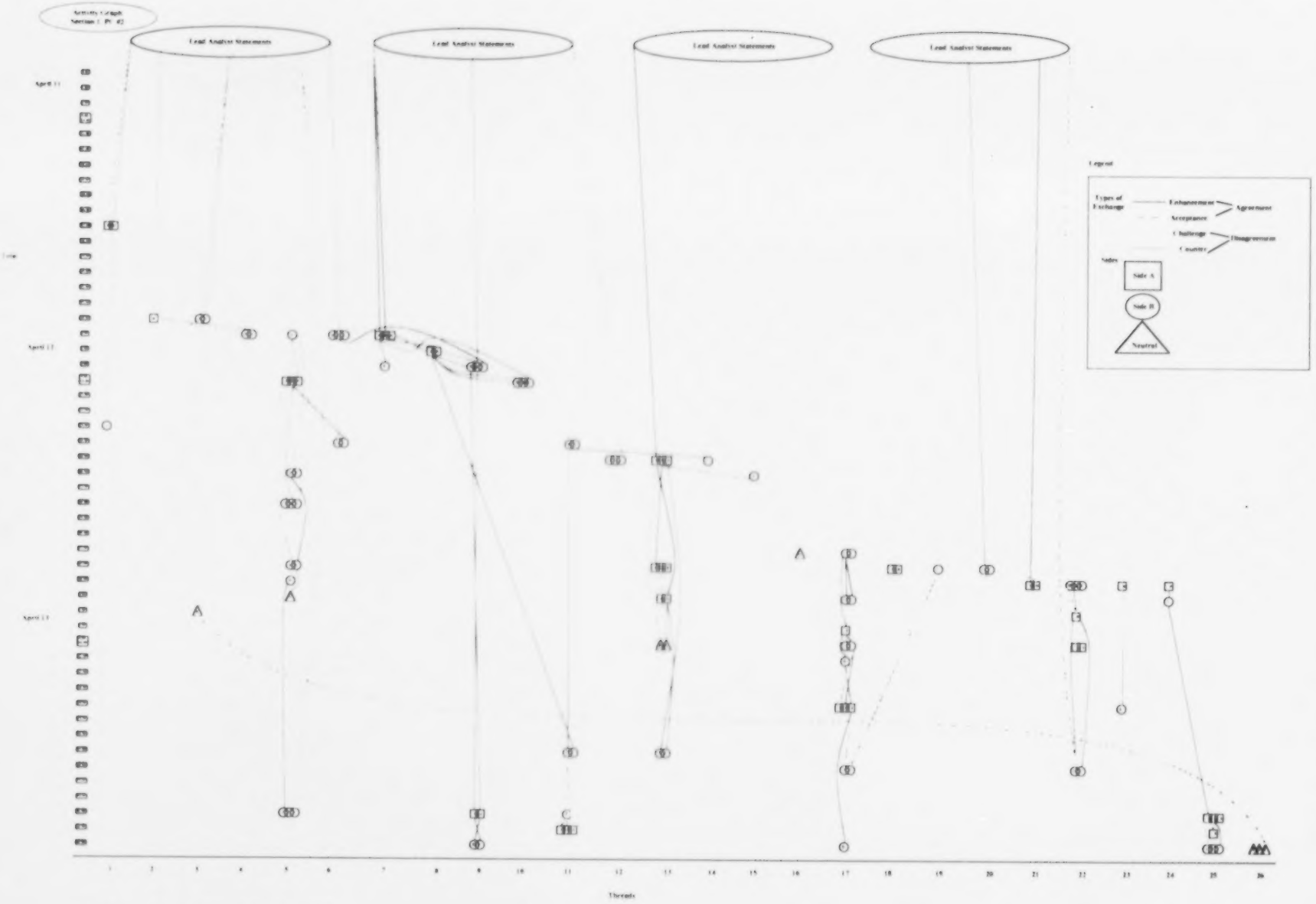

Figure 2. Activity graph for Section 1, policy challenge #2
Green lines indicate agreement with the previous point it's referring to, red lines represent disagreement. Solid lines indicate adding a new element, dashed lines indicate no new element added. Squares indicate support for Option A, circles indicate support for Option B, triangles indicate neutral

BEST COPY AVAILABLE

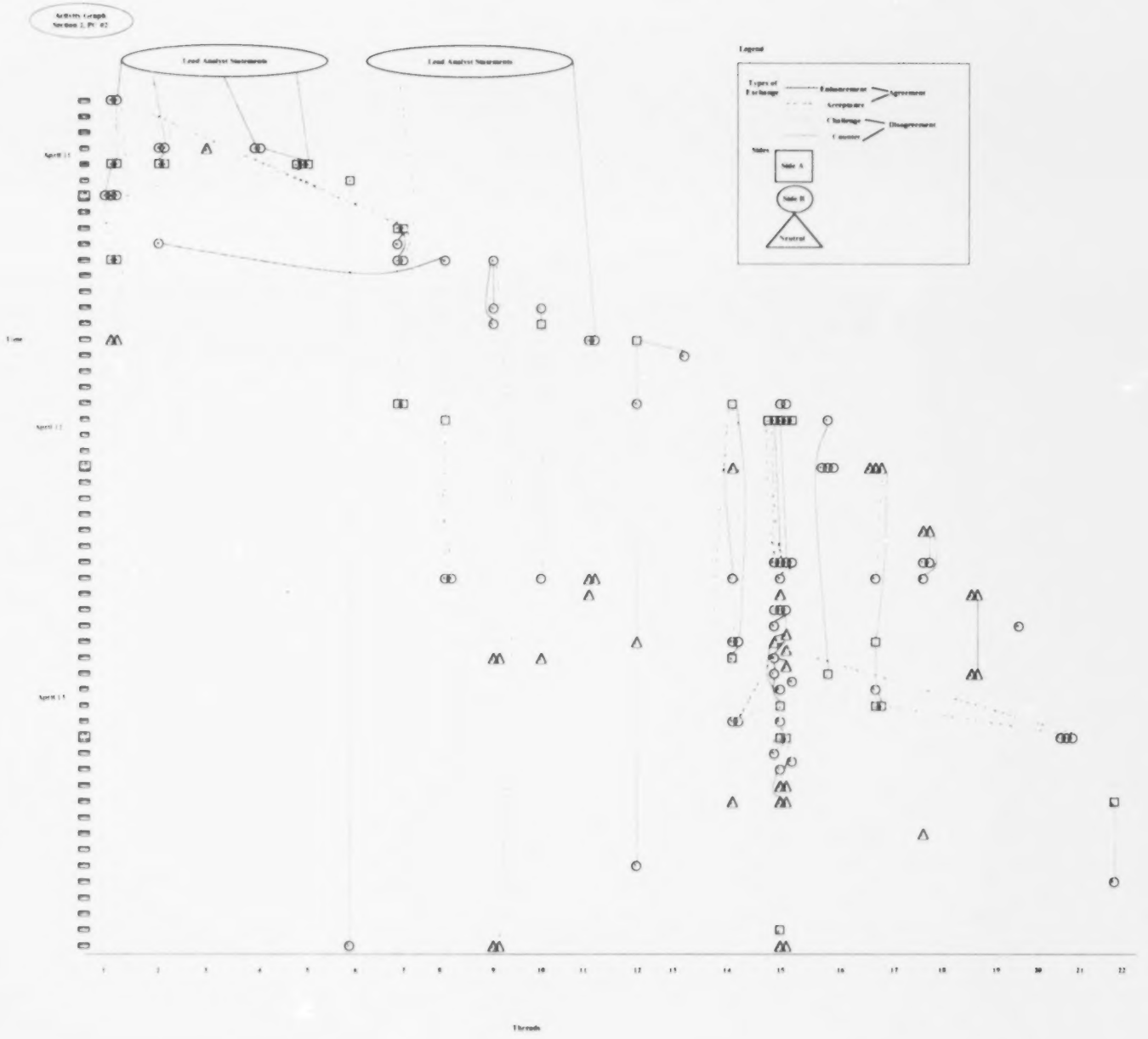

Figure 3. Activity graph for Section 2, policy challenge #2.
Green lines indicate agreement with the previous point it's referring to, red lines represent disagreement. Solid lines indicate adding a new element, dashed lines indicate no new element added. Squares indicate support for Option A, circles indicate support for Option B, triangles indicate neutral.

BEST COPY AVAILABLE

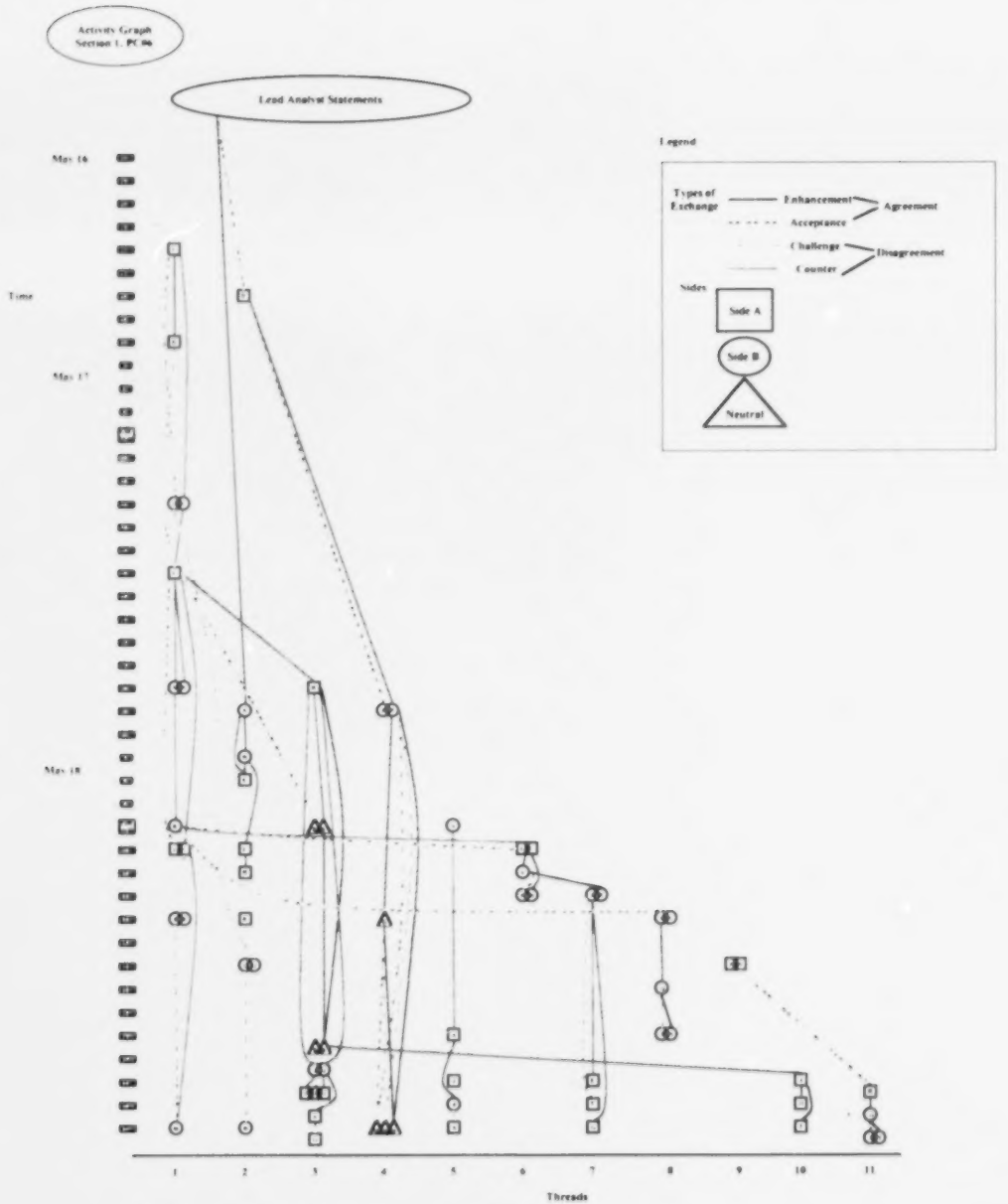

Figure 4. Activity graph for Section 1, policy challenge #6.
Green lines indicate agreement with the previous point it's referring to; red lines represent disagreement. Solid lines indicate adding a new element; dashed lines indicate no new element added. Squares indicate support for Option A; circles indicate support for Option B; triangles indicate neutral.

BEST COPY AVAILABLE

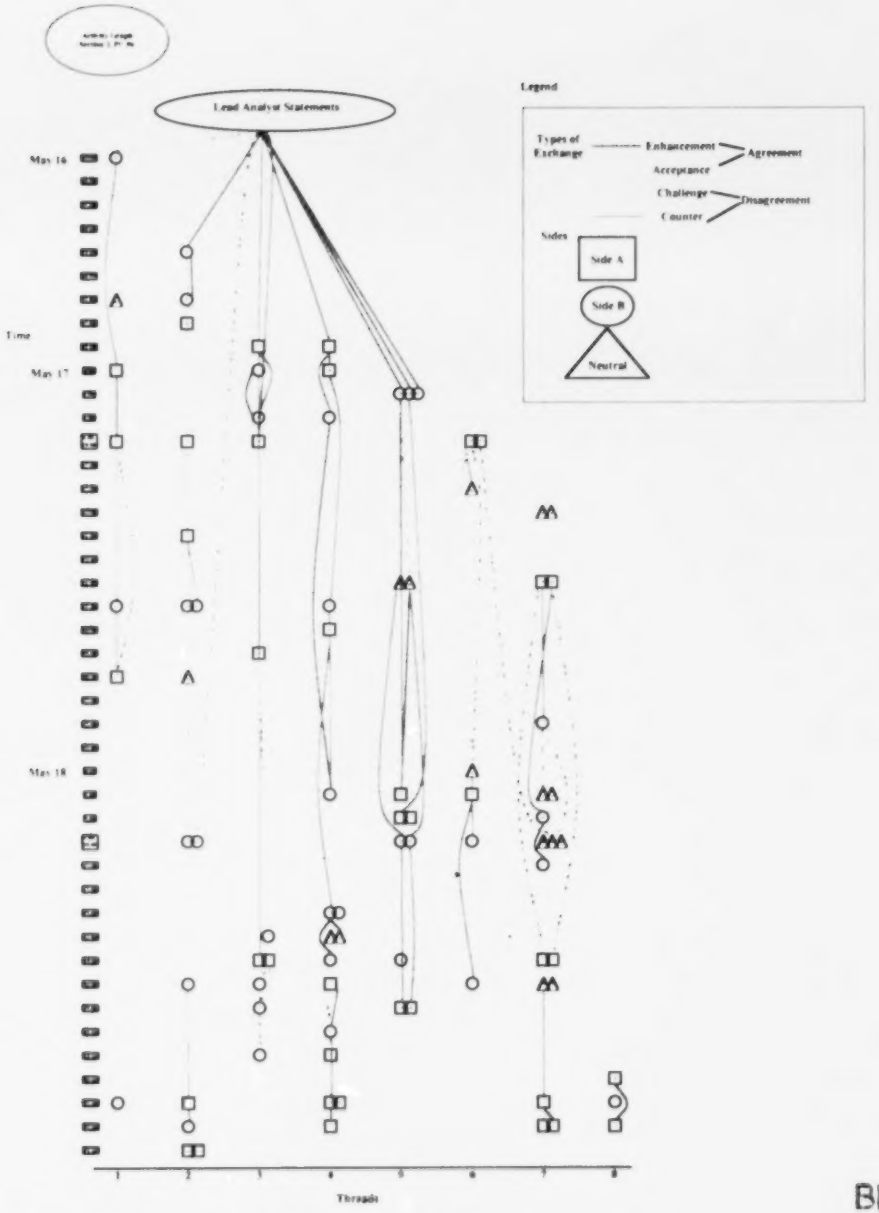

Figure 5. Activity graph for Section 2, policy challenge #6.
Green lines indicate agreement with the previous point it's referring to; red lines represent disagreement. Solid lines indicate adding a new element; dashed lines indicate no new element added. Squares indicate support for Option A; circles indicate support for Option B; triangles indicate neutral

BEST COPY AVAILABLE

Students' Experiences

During the quarter that students were participating with the policy challenges, we conducted surveys and interviews with the students to better understand how they experienced the assignment. We found that most students in the class enjoyed the opportunity to debate with and learn from each other online: "I think it's a cool idea to have a debate." Although some students were uncomfortable that the nature of the assignment was different from what they had typically experienced—students had online problem sets in previous quarters—many students enjoyed the sometimes antagonistic roles: "There's always a hole in someone's argument. You can always argue another way; because if someone thinks some way, there's always a counter-argument to it. And the trick is finding what that counter-argument is." More than 90% of the students rated the peer-to-peer debate as "especially" helpful for their learning. Most of the students reported that the policy challenges taught them how to think about issues in depth (92%) and how to improve their persuasive argumentation (70%).

More than 70% of the students reported typically reading all of the students' forum postings. "I want to do one posting, and then see what other people say." Students indicated that they learned as they read the postings of their peers: "I also think it's cool how it evolves to people thinking of new things, and you learn more by reading other people's comments."

Students differed in how open they were to the arguments made by their peers in the debate forum. Some students felt that they had to defend their position regardless of the evidence other students produced against it: "Sometimes I'm not really too open-minded about what other people have to say, because ... this is what I think, and I don't want them to win out and make me change my mind." Many students readily acknowledge the contribution of their peers to their own thinking; 74% of the students reported that they at least occasionally changed their positions on the policy challenge based on arguments from their peers. "I try to acknowledge when an argument against my position is good. And if I can't think of anything against it, then I guess maybe I should change mine," "This is a stronger argument now, so I'm going to vote for it. It's not like sticking with anything I did originally."

Many students drew distinctions between the experiences of posting early on the forum and posting later. According to one, "I was concerned about making the first posting, because absolutely everyone in section will read that. And it kind of sets the tone. I want to make sure I get it right. Whereas if I post five minutes before deadline, no one reads it anyway." Of the students who showed a clear preference between posting early and posting late on the debate forum, twice as many students preferred posting early to posting late. Table 5 displays examples of explanations that students provided.

BEST COPY AVAILABLE

Table 5
Student Responses to Posting Time Preferences.

Early is better

- "It is so much easier to post on the debate earlier. You can make general comments and expand on them to your heart's content. If you post later everything has already been said."
- "I prefer posting early on the debate forum because the issues are not all discussed already. When I post late all of the topics have been dealt with and it is hard to come up with something to post."
- "Since I personally am quite time-constrained during the week, I felt that debates often were flying by. The few times I got to post early I found myself more curious to see what people would post in response."

Later is better

- "late was better, so you could read everyone's arguments and respond."
- "more things to argue about since people have already posted their ideas"
- "late was better. by then someone had made a stupid argument that you could tear apart."
- "I liked posting late because there was more fuel for debate and I felt a little more informed then. I felt stupid not knowing a lot about any of the topics and having to make general debate questions."
- "I always ended up doing it later because of my time commitments, but I think it would have been easier to do it earlier. I feel that I learned more about argumentation, though, because of the challenge related to coming up with a cohesive argument that was not repeating what was already said.

Both better

- "I preferred a little of both. Posting early allowed you to focus on the LA response. Posting late allowed you to tackle someone else's argument. The latter is a bit more fun."
- "Both; I prefer posting early to give out my initial ideas and then posting later in response to other's ideas and challenging others' response to my stance as well.

Middle is better:

- "I prefer posting in the middle so that I can respond/counterargue people, but not late enough that everything has already been said"
- "In the middle is best: there are enough people to respond to, but every argument hasn't already been brought up, and worn out."

Although postings were the starting unit for the current analyses, students often interacted with the online debate forum without posting. In the two sections, 77% and 90% of the students reported visiting the debate forum without posting at least occasionally, and some reported doing this quite often. Unfortunately, we were unable to capture this "lurking" behavior, and explore its correlates. Moreover, we were unable to capture the visits to the debate forum after the deadline had passed, to see how often and which types of students revisited the issues, for example when studying for exams.

Discussion

Many aspects of the policy challenge assignment were successful. Although not a classroom assignment that students were immediately comfortable or even familiar with, students learned to appreciate the opportunity to challenge each others' ideas in an online environment and take and defend a stand on topics that do not have correct answers. According to one student, "The cool thing is, you can go on the web, or you're reading the paper, and it totally relates." Features of the assignment, including using authentic situations that present difficult policy choices, publicly taking a policy position, and having students synthesize the opening arguments, apparently increased the students' commitment to the debate forum, compelling them to build on each other's posts and post well more than the required amount, and more than has been typically seen (e.g., Hara et al., 2000; Howell-Richardson & Mellar, 1996; Pena-Shaff et al., 2001).

Markers of this level of participation suggest that, over time, discussions became more sustained, more focused, and better elaborated, and that individual posts became more focused (fewer points per post). Section level trends are especially apparent when comparing the activity graphs from early to late in the quarter.

However, using these indicators to make inference to student learning showed mixed results. In each section, half again as many students improved the quality of their postings (n = 12) compared to those whose posting qualities declined (n = 8). However, the increase in students' average |Exchange| was statistically significant in only one section. This mixed result is difficult to interpret: it may be subject to measurement errors (e.g., reliability of change score) or mediating variables, such as individual differences in how students benefit from such an assignment (Hartley & Bendixen, 2001; Pena-Shaff et al., 2001). Our measurements of learning styles did not show differences at the section level: average scores on the two learning styles variables did not differ significantly between sections, and therefore do not help us to understand why the improvement in posting quality does not replicate across sections. Similarly, the sections did not significantly differ in terms of the students' visiting the forum without posting.

That students experience this assignment--and the teaching in this course as a whole-- differently was well recognized by the faculty member: "There's a population of students that are excited and stimulated by the opportunity to think about problems, to have the kind of discourse that we provide for them, both in the discussion sections and the Internet, that like the idea of a course that is relatively [low] in memorizable content, but pretty high in the challenges to synthesize....But, I finally concluded that there are students who just aren't ready for that, don't like it, feel pressed by it, feel frustrated by it." One student expressed frustration as follows: "We never did policy before. So there haven't really been other sides, so much. It's just been...well, research shows this. But research doesn't show whether tradable permits are better than something else." Another student commented that participating on the debate forum helped him to better understand the issues: "I think more about it. Whereas when I have a textbook, I'm just trying to learn what's in the textbook."

The most notable shortcoming of this evaluation study is the lack of independent measure of student learning in terms of improving argumentation. While not intended for individual assessments, the coding schemes and activity graphs developed here give a sense of section-level behavior. The graphic displays are particularly helpful for conveying the nature of the online forum activity and provide at-a-glance representations of the progress of a given debate and of the differences between sections or within a section across time. They highlight the times that the students tend to post, the postings that are unconnected and don't relate to previous ideas, and the within-thread and cross-thread referrals.

The benefits of graphically displaying the online debate activity have just begun to be realized. Future directions might include characterizations of interaction patterns, similar to Hara et al. (2000; "starter-centered," "synergistic," etc.) and how the structural patterns in the interactions vary as a function of the design of the assignment (e.g., Howell-Richardson & Mellar, 1996) and features of the online tool (e.g., explicit commenting and referencing).

One could envision real-time mapping of the interactivity on the debate forum for the use of the participants and/or a moderator. For example, students might be required to comment on the evolution of the debate, to help them to recognize which types of posts are helpful for stimulating and furthering the debate, and which ones aren't. Moreover, incorporating this type of analysis back into the design of an assignment would ensure that students realized the full benefits that visiting and analyzing the debate forum offers. As it stands, there was nothing in the assignment preventing students from posting early in the debate and never returning to read and benefit from the ensuing posts of their peers. Moreover, much of the evaluation reported here hinges on actual posting behavior. Future studies should take into account benefits that students might receive from visiting the online debate forum without posting.

BEST COPY AVAILABLE

References

Boyer Commission on Educating Undergraduates in the Research University (1998). Reinventing undergraduate education: A blueprint for America's research universities. Princeton, N.J.: Carnegie Foundation for the Advancement of Teaching.

Bruffee, K. A. (1999). Collaborative learning. Higher education, interdependence, and the authority of knowledge (2nd ed.). Baltimore: Johns Hopkins University Press.

Craig, D., ul-Haq, S., Khan, S., Zimring, C., Kehoe, C., Rick, J., & Guzdial, M. (2000, June). Using an unstructured collaboration tool to support peer interaction in large college classes. In the Proceedings of the International Conference of the Learning Sciences. Ann Arbor: The University of Michigan.

Engel, C., & Schaeffer, E. (2001). Learning to persuade and persuading to learn: Design and evaluation of an online debate forum for large lecture classes. Poster presented at the European CSCL Conference, 2001, Maastricht.

Friedlander, L., & Kerns, C. (June 1998). Transforming the large lecture course. Syllabus Magazine.

Hara, N., Bonk, C., & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. Instructional Science, 28(2), 115-152.

Hartley, K., & Bendixen, L.D. (2001) Educational research in the internet age: Examining the role of individual characteristics. Educational Researcher, 22-26.

Henri, F. (1992). Computer conferencing and content analysis. In Kaye, A. R. (Ed). Collaborative Learning through Computer Conferencing: The Najaden Papers. Berlin: Springer-Verlag.

Howell-Richardson, C., & Mellar, H. (1996). A methodology for the analysis of patterns of participation within computer mediated communication courses. Instructional Science, 24(1), 47 – 69.

Jermann, P., & Dillenbourg, P. (1999). An analysis of learner arguments in a collective learning environment. In C. M. Hoadley and J. Roschelle (Eds.), Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference. Palo Alto, CA: Stanford University.

Lunsford, A. A., & Ruszkiewicz, J. J. (1999). Everything's an argument. Boston: Bedford/St. Martin's.

Martunnen, M. (1998). Electronic mail as a forum for argumentative interaction in higher education studies. Educational Computing Research, 18(4), 387 - 406.

Pena-Shaff, J., Martin, W., & Gay, G. (2001). An epistemological framework for analyzing student interactions in computer-mediated communication environments. Journal of Interactive Learning Research, 12(1), 41 - 68.

Policy challenges: Overview, expectations, and assessment. (2000). Professor Don Kennedy. Course handout, Human Biology 4B.

Roschelle, J. (1992). Learning by collaborating. Convergent conceptual change. Journal of the Learning Sciences, 2(3), 235 - 276.

Schaeffer, E., McGrady, J., Gallardo, S., Michalchik, V., Martin, N., Birks, H., & Nash, J. (1999). Stanford Learning Lab and Human Biology 4B: Evaluation Report. Online Policy Challenges and Group Websites. Year One. Stanford, CA: Stanford Learning Lab. Available at: http://www.learninglab.stanford.edu/projects/humbio.

Simons, H. W. (1976). Persuasion: Understanding, practice and analysis. Reading, MA: Addison-Wesley.

Twigg, C. A. (1999). Improving learning and reducing costs: Redesigning large-enrollment courses. The Pew Learning and Technology Program, Center for Academic Transformation. Troy, N.Y.

Van Eemeren, F., Grootendorst, R. & Kruiger, T. (1987). Handbook of argumentation theory. Dordrecht-Holland: Foris Publications.

Veerman, A., Andriessen, J., & Kanselaar, G. (1999). Collaborative learning through computer-mediated argumentation. In C. M. Hoadley and J. Roschelle (Eds.), Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference (pp. 260-264). Palo Alto, CA: Stanford University.

Voss, J. F., & Means, M. L. (1991). Learning to reason via instruction in argumentation. Pittsburgh: Learning Research and Development Center, University of Pittsburgh.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Online Debate to Encourage Peer Interactions in the Large Lecture Setting: Coding and Analysis of Forum Activity.	
Author(s): Evonne L. Schaeffer, Jennifer A. McGrady, Tina Bhargava, Claudia Engel	
Corporate Source:	Publication Date: April 2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1 ↑ [X]

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A ↑ [ ]

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B ↑ [ ]

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Evonne Schaeffer	Printed Name/Position/Title: Evonne Schaeffer, Asst Director	
Organization/Address: Stanford University	Telephone: (650) 855-5760	FAX:
135 Glenwood Ave · Woodside · CA · 94062	E-Mail Address: evonne@stanford.edu	Date: 4/3/02

(over)